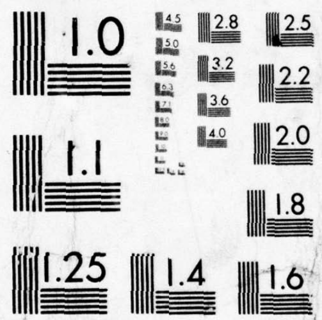
AD-A031 968 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
DISCRETE LEAST SQUARES APPROXIMATIONS FOR ORDINARY DIFFERENTIAL--ETC(U)
JUL 76 U ASCHER DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1654 NL

1 OF 1
AD
A031968

END
DATE
FILMED
1 - 77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A031968

MRC Technical Summary Report # 1654

DISCRETE LEAST SQUARES APPROXIMATIONS FOR ORDINARY DIFFERENTIAL EQUATIONS

Uri Ascher

See 1473

UNIVERSITY OF WISCONSIN

MATHEMATICS RESEARCH CENTER

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

DDC
RECEIVED
1976
C

July 1976

(Received June 16, 1976)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

DISCRETE LEAST SQUARES APPROXIMATIONS FOR ORDINARY DIFFERENTIAL EQUATIONS

Uri Ascher

Technical Summary Report #1654
July 1976

ABSTRACT

The application of the least squares method, using C^q piecewise polynomials of order $k + m$, $k \geq m$, $q \geq m$, for obtaining approximations to an isolated solution of a nonlinear m-th order ordinary differential equation, involves integrals which in practice need to be discretized. Using for this latter purpose the k-point Gaussian quadrature rule in each subinterval, the discrete least squares schemes obtained are close to collocation, on the same points, by piecewise polynomials from C^{m-1}.

We prove here that under smoothness assumptions similar to those made by de Boor and Swartz for the collocation procedure, i.e. that the solution be in C^{m+2k}, an optimal global rate of convergence $O(|\Delta|^{k+m})$ is obtained in the uniform norm for the discrete least squares schemes, provided that the partitions Δ are quasiuniform. In addition, a super-convergence rate of $O(|\Delta|^{2k})$ is obtained at the knots for those derivatives ℓ which satisfy $0 \leq \ell \leq 2(m-1) - q$.

AMS (MOS) Subject Classification: 65L10

Key Words: Discrete least squares, Gaussian points, Collocation

Work Unit Number 7 (Numerical Analysis)

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

DISCRETE LEAST SQUARES APPROXIMATIONS FOR ORDINARY DIFFERENTIAL EQUATIONS

Uri Ascher

1. Introduction

One of the well-known methods for approximately solving differential equations using piecewise polynomial functions, is the least squares method. It has been rigorously treated for linear elliptic partial differential equations by Bramble and Schatz [8] and others, and "optimal" global rates of convergence were established in the L_2-norm.

The application of the method involves integrals, which in practice should almost always be replaced by quadrature (or cubature) sums. A related question is then, how this discretization should be performed in order to retain the high order accuracy of the continuous method without using too much computing work.

In this paper we deal with ordinary differential equations. For the discretization of the integrals involved in the least squares method here, we use Gaussian quadrature rules, an obvious choice. The resulting schemes are then close, in a sense, to the collocation method of de Boor - Swartz [5]. In fact, the functions involved are evaluated on exactly the same Gaussian points, the difference between

---

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

the methods being in the degree of continuity of the approximation spaces: in [5] the piecewise polynomial space is chosen so that there are enough degrees of freedom to satisfy the differential equation exactly on the Gaussian points, whereas here the degree of continuity of the approximating functions at the knots (joints) is raised, tying up some free parameters and resulting in an overdetermined set of equations which is solved in the weighted least squares sense.

We shall prove here, under the same assumptions of smoothness on the problem as in [5], that an optimal global rate of convergence is achieved in the uniform norm, when using the discrete least squares procedure outlined above. In addition, a superconvergence result similar to that in [5], but for less derivatives, will be proven. These results extend those of [19], which verified a conjecture made in [18].

In order to make these statements more precise, we introduce some notation at this point.

Consider the linear problem of finding $u(\cdot)$ which satisfies

$$(1.1) \qquad Lu(t) = D^m u(t) - \sum_{\ell=0}^{m-1} a_\ell(t) D^\ell u(t) = f(t), \quad a \le t \le b, \; m > 0$$

and

$$(1.2) \qquad \beta_\ell(u) = 0, \quad \ell = 1, \ldots, m$$

where β_ℓ are m linearly independent side conditions. We shall take, e.g.,

$$(1.3) \qquad \beta_\ell \in \operatorname{span}\{\delta_a, \ldots, \delta_a^{(m-1)}, \delta_b, \ldots, \delta_b^{(m-1)}\}, \quad \ell = 1, \ldots, m$$

i.e., each β_ℓ is a linear combination of evaluations at a and/or b of derivatives of order $\leq m - 1$.

Let $a_\ell \in C^\ell[a, b]$, $\ell = 0, \ldots, m - 1$, and assume that (1.1) - (1.2) is uniquely solvable, i.e. that Green's function $G(t, s)$ for this problem exists, whence

$$D^\ell u(t) = \int_a^b G_\ell(t, s) f(s) ds \tag{1.4}$$

where

$$G_\ell(t, s) = \left(\frac{\partial}{\partial t}\right)^\ell G(t, s), \quad \ell = 0, \ldots, m - 1 . \tag{1.5}$$

We approximate u by an element of the piecewise polynomial spline space $S_r^q(\Delta)$, defined as follows. Let Δ be a partition of $I = [a, b]$

$$\begin{cases} \Delta : a = t_0 < t_1 < \ldots < t_N = b \\ I_i = (t_{i-1}, t_i), \; h_i = t_i - t_{i-1}, \quad i = 1, \ldots, N \\ |\Delta| = h = \max\limits_{1 \leq i \leq N} h_i . \end{cases} \tag{1.6}$$

Let $\mathcal{P}_r(E)$ denote the class of polynomials of degree $< r$ on $E \subset I$. Define, for $0 \leq q \leq r - 2$,

$$S = S_r^q(\Delta) = \{v \in C^q(I); v|_{I_i} \in \mathcal{P}_r(I_i), \quad i = 1, \ldots, N\} . \tag{1.7}$$

Let now $r = k + m$, $k \geq m$, $q \geq m$. The approximate solution will be sought in the space

$$\mathring{S} = \mathring{S}_{k+m}^q(\Delta) = S_{k+m}^q(\Delta) \cap \{v; \beta_\ell(v) = 0, \quad \ell = 1, \ldots, m\} . \tag{1.8}$$

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL

The discrete least squares approximation will be defined on Gaussian points: with

$$-1 < \rho_1 < \cdots < \rho_k < 1$$

the zeros of Legendre polynomial of degree k, define the Gaussian points as

(1.9) $$\tau_{ij} = \frac{1}{2}(t_{i-1} + t_i + \rho_j h_i), \quad i = 1, \ldots, N, \; j = 1, \ldots, k .$$

Also, let $w_1, \ldots, w_k$ be the quadrature weights associated with $\rho_1, \ldots, \rho_k$.

Throughout, let K, C be generic constants, independent of h.

Smoothness assumptions: The following smoothness assumptions will be in force throughout: For some integer n with $0 < n \leq k$, we have

(1.10) $$\begin{aligned} &f, a_\ell \in C^{(n+k)}(I), \quad \ell = 0, \ldots, m-1 \\ &u \in C^{(n+k+m)}(I) . \end{aligned}$$

It was shown in [5] that if we collocate on $\{\tau_{ij}\}$ defined in (1.9) by $v_c \in \overset{\circ}{S}{}^{m-1}_{k+m}(\Delta)$, we obtain, under these smoothness assumptions, a global accuracy of

(1.11) $$\|D^\ell(u - v_c)\|_{L_\infty(I)} \leq Ch^{k+\min(n, m-\ell)}, \quad \ell = 0, \ldots, m$$

while at the knots

(1.12) $$|D^\ell(u - v_c)(t_i)| \leq Ch^{k+n}, \quad \ell = 0, \ldots, m-1 .$$

If we now try to collocate on $\{\tau_{ij}\}$ by "smoother" spaces $S^q_{k+m}(\Delta)$, $q > m-1$, we obtain more conditions (linear equations) than

free parameters ("unknowns"), i.e., an overdetermined set of linear equations. We weigh the equation corresponding to τ_{ij} by $h_i w_j$ and solve the resulting system via least squares approximation. This can also be viewed as a discretization of the integrals resulting from the application of the least squares method to solve (1.1)-(1.2), by the Gaussian quadrature rule. Precisely, define for $z, v \in L_2(I)$

$$(1.13)\quad \begin{cases} (z,v)_i = \int_{I_i} zv\,dt, \quad 1 \le i \le N; \qquad (z,v) = \sum_{i=1}^{N} (z,v)_i = \int_I zv\,dt \\ \|z\|^2_{L_2(I_i)} = (z,z)_i, \quad 1 \le i \le N; \qquad \|z\|_2^2 = \sum_{i=1}^{N} (z,z)_i = (z,z) \end{cases}$$

$$(1.14)\quad \begin{cases} \langle z,v\rangle_i = h_i \sum_{j=1}^{k} w_j z(\tau_{ij}) v(\tau_{ij}), \quad 1 \le i \le N; \quad \langle z,v\rangle = \sum_{i=1}^{N} \langle z,v\rangle_i \\ |z|^2_{\ell_2(I_i)} = \langle z,z\rangle_i, \quad 1 \le i \le N; \qquad |z|_2^2 = \langle z,z\rangle . \end{cases}$$

(Then, for $z \cdot v \in \mathcal{P}_{2k}(I_i)$ we have $\langle z,v\rangle_i = (z,v)_i$, and if this holds for $i = 1, \ldots, N$, we have $\langle z,v\rangle = (z,v)$). For a fixed $q, q \ge m$, the approximate solution provided by the discrete method, $v^* \in \mathring{S}$, is determined by

$$(1.15)\qquad \min_{v \in \mathring{S}} |Lv - f|_2^2 .$$

The solution provided by the "continuous" method, $w^* \in \mathring{S}$, is determined by

$$(1.16)\qquad \min_{w \in \mathring{S}} \|Lw - f\|_2^2 .$$

We shall prove the following results for both v^* and w^*:

Assume that the partition Δ is quasiuniform, i.e. $\exists C$ such that

$$h/\min_{1 \le i \le N} h_i \le C \tag{1.17}$$

and C does not depend on h. Then, under the smoothness assumptions (1.10), an optimal global error estimate is obtained in the uniform norm,

$$\| D^{\ell}(u - v^*) \|_{L_\infty(I)} \le O(h^{k+\min(n, m-\ell)}), \quad \ell = 0, \ldots, m . \tag{1.18}$$

Also, at the knots, with $\mu := 2(m - 1) - q \ge 0$,

$$| D^{\ell}(u - v^*)(t_i) | \le O(h^{k+n}), \quad \ell = 0, \ldots, \mu . \tag{1.19}$$

If $\mu < 0$ then no superconvergence will occur at the knots.

Note that, with $q = m - 1$, $\mu = m - 1$ and (1.19) reduces to (1.12). These superconvergence results are similar to those obtained in [11] for the H^1-Galerkin procedure for the heat equation. See also [12].

An outline of the remainder of the paper follows. In Section 2, we prove that the numerical schemes (1.15) are well-defined for all small enough h. We then proceed to estimate distances between discrete and continuous inner products. In Section 3, we treat the special case $L \equiv D^m$. This serves both for illustration and for obtaining, later on, the desired results for the general linear case. In Section 4, optimal global results are proven in the L_1-norm and in Section 5, results (1.18)-(1.19) are obtained for the general linear case.

Section 6 treats the "continuous", or "integral", least squares approximation (1.16), and the nonlinear case. For the nonlinear problem,

a Newton process of linearization and iteration is applied. The linearized problem is solved as outlined above. Thus, roughly speaking, the method would apply (as in [5]), provided that the nonlinear problem is "decently linearizable" around an isolated solution and the starting guess is close enough to that solution. In Section 7, numerical examples are given, demonstrating the claims above and comparing the discrete least squares schemes to the collocation scheme of de Boor-Swartz [5]. The collocation scheme is found to be better (not by much, though) in every respect, except that the dimension of the approximation space gets larger for collocation, resulting in some additional storage requirements for the coefficients of the computed solution. This, however, is not a serious factor here, and we are led to conclude that collocation is generally superior to discrete least squares for ordinary differential equations.

2. Discrete and continuous inner products

Let

$$e = v^* - u \tag{2.1}$$

and

$$R = Lv^* - f = Le \tag{2.2}$$

where $v^* \in \overset{\circ}{S}{}^q_{k+m}(\Delta)$ is the solution of (1.15). By the Green's function representation we have, for $0 \leq \ell \leq m - 1$,

$$D^\ell e(t) = \int_a^b G_\ell(t, s)R(s)ds \equiv (G_\ell(t), R) = [(R, G_\ell(t)) - \langle R, G_\ell(t)\rangle] + \langle R, G_\ell(t)\rangle . \tag{2.3}$$

Our main aim in this section is to estimate the first term on the right hand side of (2.3). Before that we shall prove that (1.15) has a unique solution v^* (cf. [19]).

Let $W^s_p(J)$ and $H^s(J)$ denote the closure of $C^\infty(\bar{J})$ in the norms

$$\|v\|_{W^s_p(J)} = \sum_{j=0}^{s} \|D^j v\|_{L_p(J)} \tag{2.4}$$

and

$$\|v\|_{H^s(J)} = \Big(\sum_{j=0}^{s} \|D^j v\|^2_{L_2(J)} \Big)^{1/2} , \tag{2.5}$$

respectively. In case that $J = I$, it will be dropped in the notation.

Define also

$$\|v\|_{H^{s,\Delta}} = \Big(\sum_{i=1}^{N} \|v\|^2_{H^s(I_i)} \Big)^{1/2} . \tag{2.6}$$

The smoothness assumptions (1.10) are sufficient so that

(i) $\exists K$ such that when $G_\ell(t,\cdot)$ is considered as an element of $C^{(n)}[a,t] \times C^{(n)}[t,b]$,

$$\|D^j G_\ell(t,\cdot)\| \le K, \quad t \in I, \quad 0 \le \ell \le m-1, \quad j = 0,\ldots,n. \tag{2.7}$$

(ii) Given $f \in H^p$, $0 \le p \le k+n$, the unique solution u for (1.1) - (1.2), assumed to exist, lies in $H^{p+m}(I)$ and

$$\|Lu\|_{H^p(J)} \le C\|u\|_{H^{p+m}(J)}, \quad J \subset I, \tag{2.8}$$

for some C depending only on L. We also have that

$$\|Lu\|_2 \ge C\|u\|_{H^m}. \tag{2.9}$$

Theorem 2.1.

The discrete least squares approximation (1.15) is well-defined, for h small enough.

Proof.

The system actually solved is the linear system of normal equations

$$A\alpha = b \tag{2.10}$$

for the coefficients α of the representation of v^* in terms of a specified basis of the approximation space $\mathring{S} = \mathring{S}^q_{k+m}(\Delta)$. The matrix A is symmetric positive semidefinite, and the solution of (2.10) will provide the unique minimum of (1.15) iff A is positive definite. This will hold, in turn, iff for any $v \in \mathring{S}$, $v \not\equiv 0$, we have $|Lv|_2 > 0$. We shall now prove this last statement, for sufficiently small h.

Let $v \in \mathring{S}$. Then $D^j v|_{I_i} = 0$ for $j \geq k+m$, $i = 1, \ldots, N$. Thus

$$\|v\|_{H^s,\Delta} = \|v\|_{H^{k+m-1},\Delta}, \quad s \geq k+m .$$

Expanding $(Lv)^2$ in each subinterval I_i in Taylor's series around, say, t_i, up to terms of order $k+n$, and using the Gaussian quadrature precision together with (2.8), we get

$$\left| \|Lv\|_2^2 - |Lv|_2^2 \right| \leq Ch^{k+n} \|v\|^2_{H^{k+n+m},\Delta} = Ch^{k+n} \|v\|^2_{H^{k+m-1},\Delta} .$$

In each subinterval, v is a polynomial. Using Markov's inequality, we get

$$\left| \|Lv\|_2^2 - |Lv|_2^2 \right| \leq Kh^{n+1} \|v\|^2_{H^m} . \tag{2.11}$$

Now, using (2.9) and (2.11),

$$|Lv|_2^2 = \|Lv\|_2^2 + [\, |Lv|_2^2 - \|Lv\|_2^2] \geq C\|v\|^2_{H^m} - Kh^{n+1}\|v\|^2_{H^m} .$$

For h small enough, we can therefore write

$$|Lv|_2^2 \geq C\|v\|^2_{H^m} . \tag{2.12}$$

Q. E. D.

We now prove two lemmas about the residual error, R of (2.2).

<u>Lemma 2.1</u>.

There exists K such that, for $h > 0$ small enough,

$$|R|_2 \leq Kh^k \tag{2.13}$$

$$\|R\|_2 \leq Kh^k . \tag{2.14}$$

Proof.

Let $\bar{v} \in \mathring{S}$ be such that

$$\|L\bar{v} - f\|_{L_\infty(I)} \leq O(h^k)$$

(e.g. Galerkin or collocation solution [10], [16], [17], [15], or an approximation to u [2]). Then, since v^* minimizes (1.15),

$$|R|_2 \leq |L\bar{v} - f|_2 \leq O(h^k) ,$$

obtaining (2.13). To obtain (2.14) note that

$$\|R\|_2 \leq \|Lv^* - L\bar{v}\|_2 + \|L\bar{v} - f\|_2 .$$

We are therefore left to estimate $L\eta$, where

$$\eta := v^* - \bar{v} \in \mathring{S} .$$

Clearly, we have by (2.13)

$$|L\eta|_2 \leq O(h^k) .$$

By (2.8) and (2.12), then,

$$\|L\eta\|_2 \leq C\|\eta\|_{H^m} \leq K|L\eta|_2 \leq O(h^k) . \qquad \text{Q.E.D.}$$

Lemma 2.2.

There exists K such that

(2.15) $$\|D^{k+j}R\|_{L_\infty(I_i)} \leq K, \quad i = 1,\ldots,N, \; j = 0,\ldots,n .$$

Proof.

We follow the proof of lemma 4.1 of [5]. First note that, by (2.14),

(2.16) $$|D^\ell e(t)| = |(R, G_\ell(t))| \leq \|R\|_2 \|G_\ell(t)\|_2 \leq O(h^k), \quad 0 \leq \ell \leq m-1 .$$

Thus $\|D^{\ell}e\|_{L_\infty(I_i)} \leq O(h^k)$, $\ell = 0,\ldots,m-1$. Now, since $R = Le$, $D^{k+j}R$ on I_i is a combination of derivatives of $a_\ell(\cdot)$ up to order $k+j$ with derivatives of e up to order $k+j+m$. It is therefore sufficient to prove the existence of a constant C such that

$$(2.17) \qquad \|D^j e\|_{L_\infty(I_i)} \leq C, \; i = 1,\ldots,N, \; j = 0,\ldots,m+k+n \,.$$

For $j < m$ we have (2.16). For $j \geq m+k$ we have $D^j v^* = 0$ on I_i and the result follows from the smoothness assumption on u. For $m \leq j \leq m+k-1$ let u_i be the Taylor expansion for u at $t = t_i$ up to terms of order $< m+k$. Then

$$\|D^j e\|_{L_\infty(I_i)} \leq \|D^j(v^* - u_i)\|_{L_\infty(I_i)} + \|D^j(u_i - u)\|_{L_\infty(I_i)} \,.$$

For the second term on the right hand side we clearly have

$$\|D^j(u_i - u)\|_{L_\infty(I_i)} \leq Kh^{k+m-j} \,,$$

while for the first term we have, by Markov's inequality (note that $v^* - u_i \in \mathcal{P}_{k+m}(I_i)$),

$$\|D^j(v^* - u_i)\|_{L_\infty(I_i)} \leq Kh_i^{m-1-j} \|D^{m-1}(v^* - u_i)\|_{L_\infty(I_i)}$$

and, since clearly

$$\|D^{m-1}(v^* - u_i)\|_{L_\infty(I_i)} \leq O(h^k) \,,$$

we obtain

$$\|D^j e\|_{L_\infty(I_i)} \leq Ch^{k+m-1-j}, \quad m \leq j \leq m+k-1 \,.$$

Q.E.D.

Next, we obtain the desired estimates for an interpolant of R.

Lemma 2.3.

Let $\bar{R}(\cdot) \in S_k^{-1}(\Delta)$ interpolate $R(\cdot)$ at the points $\{\tau_{ij}\}$ of (1.9). Then

(a) $\bar{R}$ exists

(b) $$\|R - \bar{R}\|_{L_\infty(I_i)} \leq O(h_i^k), \qquad 1 \leq i \leq N .$$

(c) $$\langle R, z\rangle_i = \langle \bar{R}, z\rangle_i, \quad i = 1, \ldots, N, \ z \in L_\infty(I) .$$

(d) Assume that

$$\|R\|_{L_\infty(I)} \leq O(h^k) . \tag{2.18}$$

Then

$$|(\bar{R}, G_\ell(t))_i - \langle \bar{R}, G_\ell(t)\rangle_i| \leq O(h^{k+M}), \qquad t \in I, \ 0 \leq \ell \leq m-1 \tag{2.19a}$$

where

$$M = \begin{cases} n+1, & t \notin I_i \\ \min(n+1, \ m-\ell), & t \in I_i . \end{cases} \tag{2.19b}$$

(Note that if $t = t_i$ then $M \equiv n+1$).

(e) Without assuming (2.18) we still have for $0 \leq \ell \leq m$ and $\psi \in W_\infty^{m-\ell}(I)$ satisfying the side conditions (1.2), that

$$|(\bar{R}, \psi) - \langle \bar{R}, \psi\rangle| \leq O(h^{k+m-\ell}) .$$

Proof.

(a), (b), (c) are trivial. We prove (d): Fix i such that $t \notin I_i$. Let $P \in S_n^{-1}(\Delta)$ approximate $G_\ell(t, \cdot)$ such that

$$\|P - G_\ell(t)\|_{L_\infty(I_j)} \leq O(h_j^n), \quad t \notin I_j .$$

(This is clearly possible, by (2.7)). Now,

$$(2.20) \quad (\bar{R}, G_\ell(t))_i - \langle \bar{R}, G_\ell(t)\rangle_i = [(\bar{R}, G_\ell(t))_i - (\bar{R}, P)_i] + [(\bar{R}, P)_i - \langle \bar{R}, P\rangle_i] + $$
$$+ [\langle \bar{R}, P\rangle_i - \langle \bar{R}, G_\ell(t)\rangle_i] .$$

For the second term on the right hand side we have $(\bar{R}, P)_i - \langle \bar{R}, P\rangle_i = 0$, since $\bar{R} \cdot P \in \mathcal{P}_{k+n}(I_i)$ and $n \leq k$. For the first term we write

$$|(\bar{R}, G_\ell(t) - P)_i| \leq \|\bar{R}\|_{L_2(I_i)} \|G_\ell(t) - P\|_{L_2(I_i)} \leq O(h^{k+n+1}) .$$

The above is true because

$$\|\bar{R}\|_{L_2(I_i)} \leq Ch_i^{1/2} \|\bar{R}\|_{L_\infty(I_i)}$$

and using (2.18) and (b) of the lemma we obtain

$$\|\bar{R}\|_{L_2(I_i)} \leq O(h^{k+1/2}) .$$

Similarly

$$\|P - G_\ell(t)\|_{L_2(I_i)} \leq O(h^{n+1/2}) .$$

The third term in (2.20) is treated exactly as the first one above:

$$|\langle \bar{R}, P - G_\ell(t)\rangle_i| \leq |\bar{R}|_{\ell_2(I_i)} |P - G_\ell(t)|_{\ell_2(I_i)} \leq O(h^{k+n+1}) .$$

When we now take i such that $t \in I_i$, we have $G_\ell(t, \cdot) \in W_\infty^{m-\ell-1}(\bar{I}_i)$ and thus $m - \ell - 1$ replaces n in the argument. The definition (2.19b) of M thus follows.

To prove (e) note that our result will not be local, since (2.18) is not assumed to hold. We proceed as above without breaking to subintervals: $P \in S^{-1}_{m-\ell}$ now approximates φ with

$$\|P - \varphi\|_{L_\infty(I)} \leq O(h^{m-\ell})$$

and

$$|(\overline{R}, \varphi - P)| \leq \|\overline{R}\|_2 \|\varphi - P\|_2 .$$

For $\overline{R}$ we note that

$$\|\overline{R}\|_2 \leq \|R\|_2 + \|R - \overline{R}\|_2 \leq O(h^{\kappa})$$

and combining this with the estimate on $\varphi - P$ we obtain the conclusion. Similarly for the term $|\langle \overline{R}, \varphi - P\rangle|$.

We are now ready to estimate the difference between the discrete and continuous inner products in (2.3).

<u>Theorem 2.2</u>.

With our smoothness assumptions and assuming (2.18) to hold, we have

(2.21) $$|(R, G_\ell(t)) - \langle R, G_\ell(t)\rangle| \leq O(h^{k+\overline{M}}), \quad \ell = 0, \ldots, m-1$$

where

(2.22) $$\overline{M} = \begin{cases} \min(n, m-\ell), & t \neq t_i, \text{ for all } i \\ n & t = t_i \text{ for some } i . \end{cases}$$

<u>Proof</u>.

Since $\overline{R}$ interpolates R at $\{\tau_{ij}\}$ we can write, for a fixed i,

(2.23) $$R(s) = \overline{R}(s) + P_i(s)[\tau_{i1}, \ldots, \tau_{ik}, s]R, \quad s \in I_i$$

with $[\tau_{i1}, \ldots, \tau_{ik}, s]R$ the k-th divided difference of R and

$$P_i(s) = \prod_{j=1}^{k} (s - \tau_{ij}) . \tag{2.24}$$

Now,

$$(R, G_\ell(t))_i - \langle R, G_\ell(t)\rangle_i = (R, G_\ell(t))_i - \langle \overline{R}, G_\ell(t)\rangle_i = [(\overline{R}, G_\ell(t))_i - \langle \overline{R}, G_\ell(t)\rangle_i] +$$

$$+ (P_i(\cdot)[\tau_{i1}, \ldots, \tau_{ik}, \cdot]R, \; G_\ell(t, \cdot))_i .$$

The first term on the right was estimated in Lemma 2.3(d). For the second term we proceed exactly as in [5, proof of Thm. 4.1, pp. 600-601], using Lemma 2.2. Now, to obtain (2.21) we sum up on i, $i = 1, \ldots, N$. Since we have $N - 1$ "regular" intervals with $M = n + 1$ and one "special" interval with $M = \min(n + 1, m - \ell)$, we obtain an overall error (2.21) with $\overline{M}$ given by (2.22) for $t \neq t_i$ all i.

When $t = t_i$ for some i, we simply do not have a "special" interval I_i such that $t \in I_i$. Thus, summing up N terms of order $O(h^{k+n+1})$ gives the order $O(h^{k+n})$. Q.E.D.

<u>Corollary 2.1</u>.

With $\psi(\cdot) \in W_\infty^{m-\ell}(I)$ replacing $G_\ell(t, \cdot)$ in Theorem 2.2 we obtain

$$|(R, \psi) - \langle R, \psi\rangle| \leq O(h^{k+m-\ell}), \qquad 0 \leq \ell \leq m. \tag{2.25}$$

(Note that (2.18) is not assumed to hold.)

<u>Proof</u>.

$$(R, \psi) - \langle R, \psi\rangle = [(R, \psi) - (\overline{R}, \psi)] + [(\overline{R}, \psi) - \langle \overline{R}, \psi\rangle] .$$

The second term on the right was satisfactorily estimated in Lemma 2.3(e). For the first term we again write

$$(R - \overline{R}, \psi)_i = (P_i(\cdot), \psi(\cdot)[\tau_{i1}, \ldots \tau_{ik}, \cdot]R)$$

P_i given by (2.24), and proceed exactly as in [5, pp. 600-601]. Note that the derivatives of ψ and R are taken to be bounded in the uniform norm, according to the assumption on and Lemma 2.2. Q.E.D.

3. The case of $L \equiv D^m$

In order to be able to use Theorem 2.2 and to proceed further, we have to bound R appropriately in the uniform norm (i.e., show that (2.18) holds). For this end and for illustrative purposes we obtain all the results first for the special operator $L \equiv D^m$.

Let u be the solution of (1.1)-(1.2). Then

$$D^m u(\cdot) = f(\cdot) + \sum_{\ell=0}^{m-1} a_\ell(\cdot) D^\ell u(\cdot) =: \hat{f}(\cdot) . \tag{3.1}$$

We have $\hat{f} \in C^{k+n}(I)$. Let $U(\cdot)$ be the solution of

$$\min_{v \in \mathring{S}^q_{k+m}(\Delta)} |D^m v - \hat{f}|_2^2 . \tag{3.2}$$

Since we have assumed (1.3), we are able to conclude that Green's function $H(t,s)$ for D^m exists:

$$D^\ell u(t) = \int_a^b H_\ell(t,s)\hat{f}(s)ds, \quad 0 \le \ell \le m-1 \tag{3.3}$$

where

$$H_\ell(t,s) = \left(\frac{\partial}{\partial t}\right)^\ell H(t,s) . \tag{3.4}$$

Let

$$\nu = U - u . \tag{3.5}$$

First we prove the equivalent of (2.18) for this special case.

Lemma 3.1.

$$\|D^m \nu\|_{L_\infty(I)} \le O(h^k) . \tag{3.6}$$

Proof.

Note that

$$D^m \nu = D^m U - \hat{f}$$

i.e., $D^m \nu$ is the quantity minimized in (3.2). Let now $\hat{P} \in S_k^{-1}(\Delta)$ interpolate $\hat{f}$ at $\{\tau_{ij}\}$, $i = 1, \ldots, N$, $j = 1, \ldots, k$. Then, for any $v \in \mathring{S}^q_{k+m}(\Delta)$,

$$|D^m v - \hat{f}|_2 = |D^m v - \hat{P}|_2 = \|D^m v - \hat{P}\|_2 \tag{3.7}$$

and

$$\|\hat{f} - \hat{P}\|_{L_\infty(I)} \leq O(h^k) . \tag{3.8}$$

By (3.7) and (3.2) we have U determined by

$$\min_{v \in \mathring{S}^q_{k+m}(\Delta)} \|D^m v - \hat{P}\|_2 .$$

But then $D^m U$ is just the L_2-projection of $\hat{P}$ into $S_k^{q-m}(\Delta)$. By [10], [4] we have

$$\|D^m U - \hat{P}\|_{L_\infty(I)} \leq O(h^k) .$$

Combining this with (3.8) we obtain (3.6). Q.E.D.

We are now ready for the general theorem for this special operator.

Theorem 3.1.

Under our smoothness assumptions (and for quasiuniform partitions) we have

$$\|D^\ell \nu\|_{L_\infty(I)} \leq O(h^{k+\min(n, m-\ell)}), \quad \ell = 0, \ldots, m \tag{3.9}$$

and, at the knots,

$$(3.10) \qquad |D^{\ell}\nu(t_i)| \leq O(h^{k+n}), \quad \ell = 0,\ldots,\mu$$

where $\mu = 2(m-1) - q$. (Note that for q sufficiently large, $\mu < 0$ and (3.10) occurs for no derivative, i.e., no superconvergence).

<u>Proof</u>.

For the pointwise error, we have

$$(3.11) \quad D^{\ell}\nu(t) = (D^m\nu(\cdot), H_{\ell}(t,\cdot)) = [(D^m\nu, H_{\ell}(t)) - \langle D^m\nu, H_{\ell}(t)\rangle] + \langle D^m\nu, H_{\ell}(t)\rangle .$$

For the first term on the right, Theorem 2.2 applies, since we have proved (3.6). For the remaining discrete term in (3.11) we have, since U solves (3.2),

$$(3.12) \quad |\langle D^m\nu, H_{\ell}(t)\rangle| = |\langle D^m\nu, H_{\ell}(t) - D^m v\rangle| \leq \|D^m\nu\|_{L_\infty(I)} |H_{\ell}(t) - D^m v|_1$$

$$\text{for all} \quad v \in \mathring{S}^q_{k+m}(\Delta)$$

where for a function $\psi \in L_\infty(I)$

$$(3.13) \qquad |\psi|_1 := \sum_{i=1}^{N} h_i \sum_{j=1}^{k} w_j |\psi(\tau_{ij})| .$$

Now, with this simple operator, $D^m v$ is just an approximant out of $S_k^{q-m}(\Delta)$ to $H_{\ell}(t,\cdot)$ and can clearly be chosen so that [2]

$$\|D^m v - H_{\ell}(t)\|_{L_\infty(I_i)} \leq O(h^M)$$

where, if $t \in I_{i_0}$,

$$M = \begin{cases} \min(n, m-\ell-1), & i_0 - k \leq i \leq i_0 + k \\ n & \text{otherwise} . \end{cases}$$

Multiplying each of these estimates by h_i and summing up on i, we obtain

$$(3.14) \qquad \inf_{v \in \mathring{S}^q_{k+m}(\Delta)} |H_\ell(t) - D^m v|_1 \leq O(h^{\min(n, m-\ell)}) .$$

Substituting this and (3.6) into (3.12), we obtain (3.9) (via (3.11)), for $0 \leq \ell \leq m - 1$. The conclusion for $\ell = m$ now follows from (3.6).

If $t = t_j$ for some j, $1 \leq j \leq N$, then we simply have here

$$H_\ell(t_j) \in S_k^{m-\ell-2}(\Delta) .$$

If $m - \ell - 2 < q - m$ then the above argument can be applied with $I_j \cup I_{j+1}$ replacing I_{i_0}, and (3.9) follows. Further, if our approximation space is not too smooth, then we can obtain more: if $0 \leq \ell \leq \mu$ then

$$\ell \leq 2(m - 1) - q \quad \text{implies} \quad m - \ell - 2 \geq q - m$$

and now we simply have, instead of (3.14),

$$\min_{v \in \mathring{S}^q_{k+m}} |H_\ell(t) - D^m v|_1 = 0 .$$

The result (3.10) follows from a similar result in Theorem 2.2. Q.E.D.

4. An error bound in the L_1-norm

In order to obtain pointwise error estimates we first obtain optimal global results in the L_1-norm (cf. [19], [1]).

Theorem 4.1.

$$\|e\|_{W_1^{\ell}(I)} \leq O(h^{k+\min(n, m-\ell)}), \quad \ell = 0,\dots,m . \tag{4.1}$$

Proof.

We apply a duality argument. Let ψ satisfy (1.2) and solve

$$L^*\psi = g \tag{4.2}$$

where $g \in W_\infty^{-\ell}(I)$, $\|g\|_{W_\infty^{-\ell}(I)} = 1$, $0 \leq \ell \leq m$. Assume, without loss of generality, that $n \geq m - \ell$. Then

$$(g, e) = (L^*\psi, e) = \int_a^b L^*\psi(t) \int_a^b G(t, s)R(s)dsdt = \int_a^b R(s) \int_a^b G(t, s)L^*\psi(t)dtds =$$

$$= \int_a^b R(s)\psi(s)ds = (R, \psi) .$$

We now make use of Corollary 2.1:

$$|(g, e)| \leq |(R, \psi) - \langle R, \psi\rangle| + |\langle R, \psi\rangle| \leq |\langle R, \psi\rangle| + O(h^{k+m-\ell}) .$$

For the remaining discrete term we have by (1.15)

$$\langle R, \psi\rangle = \langle R, -Lv + \psi\rangle, \quad v \in \mathring{S}_{k+m}^q(\Delta) . \tag{4.3}$$

Taking v to be, e.g., the Galerkin or collocation solution to

$$\begin{cases} L\phi = \psi \\ \phi \quad \text{satisfies (1.2)}, \end{cases}$$

we have

$$|Lv - \psi|_2 \leq O(h^{m-\ell}) \| D^{m-\ell} \psi \|_{L_\infty(I)}$$

and

$$|\langle R, \psi \rangle| \leq |R|_2 |Lv - \psi|_2 \leq O(h^{k+m-\ell}) .$$

Thus,

$$\|e\|_{W_1^\ell(I)} = \sup_g |(g, e)| \leq O(h^{k+m-\ell}) . \qquad \text{Q.E.D.}$$

<u>Corollary 4.1.</u>

$$\|D^\ell e\|_{L_1(I)} \leq O(h^{k+\min(n, m-\ell)}), \qquad \ell = 0, \ldots, m . \tag{4.4}$$

<u>Remark</u>.

Note that the quasiuniformity assumption (1.17) was not used here. It is used in obtaining a bound similar to (4.4) in the L_∞-norm.

5. L_∞ estimates and superconvergence results for a general linear operator

Here we extend the results obtained above for ν, to e. First we establish (2.18). Let

$$\eta = v^* - U . \tag{5.1}$$

Then $e = \nu + \eta$ and we are left to estimate η. Note that $\eta \in \mathring{S}^q_{k+m}(\Delta)$.

Lemma 5.1.

(a)

$$\|D^\ell \eta\|_{L_1(I)} \le O(h^{k+\min(n,m-\ell)}), \qquad 0 \le \ell \le m . \tag{5.2}$$

(b) If

$$|L\eta|_2^2 \le O(h^{2k+1}) \tag{5.3}$$

then

$$\|R\|_{L_\infty(I)} \le O(h^k) . \tag{5.4}$$

Proof.

(a) follows immediately from (3.9) and (4.4). To prove (b) assume (5.3) to hold. Similarly to (2.16), we have

$$\|D^\ell \eta\|_{L_\infty(I)} \le O(h^k), \qquad \ell = 0,\dots,m-1 . \tag{5.5}$$

Now, we have assumed

$$|L\eta|_2^2 = \sum_{i=1}^N h_i \sum_{j=1}^k w_j [L\eta(\tau_{ij})]^2 \le O(h^{2k+1}) .$$

Under quasiuniformity, then, for each i

$$\sum_{j=1}^k w_j [L\eta(\tau_{ij})]^2 \le O(h^{2k})$$

$$(5.6) \qquad |L\eta(\tau_{ij})| \le O(h^k), \qquad i = 1, \ldots, N, \; j = 1, \ldots, k\,.$$

But, by (5.5) this means

$$(5.7) \qquad |D^m \eta(\tau_{ij})| \le O(h^k), \quad i = 1, \ldots, N, \; j = 1, \ldots, k\,.$$

Fix i. On I_i, $D^m \eta \in \mathcal{P}_k(I_i)$ and the equivalence of norms on a finite dimensional space yields

$$\|D^m \eta\|_{L_\infty(I_i)} \le O(h^k), \qquad i = 1, \ldots, N\,.$$

Combining this with (5.5), we obtain

$$\|L\eta\|_{L_\infty(I)} \le O(h^k)$$

and (5.4) follows, since $R = Le = L\nu + L\eta$. Q.E.D.

Lemma 5.2.

Under our assumptions (of smoothness and quasiuniform mesh)

$$(5.8) \qquad \|R\|_{L_\infty(I)} \le O(h^k)\,.$$

Proof.

By the previous lemma we are left to prove (5.3). Since $\eta \in \mathring{S}^q_{k+m}(\Delta)$ we have

$$\langle R, L\eta\rangle = 0$$

i.e.,

$$0 = \langle L\nu + L\eta, L\eta\rangle = |L\eta|_2^2 + \langle L\nu, L\eta\rangle\,.$$

$$(5.9) \qquad |L\eta|_2^2 = -\langle L\nu, L\eta\rangle = -\langle D^m\nu, D^m\eta\rangle + \langle D^m\eta, \sum_{\ell=0}^{m-1} a_\ell D^\ell \nu\rangle +$$

$$+ \langle \sum_{\ell=0}^{m-1} a_\ell D^\ell \eta, D^m \nu\rangle - \langle \sum_{\ell=0}^{m-1} a_\ell D^\ell \nu, \sum_{\ell=0}^{m-1} a_\ell D^\ell \eta\rangle\,.$$

Now, by definition of ν, $\langle D^m \nu, D^m \eta \rangle = 0$. The other terms on the right hand side of (5.9) are of lower order. We would like to show that each of the discrete inner products appearing in (5.9) is $\leq O(h^{2k+1})$. This is immediate for all terms except for $\langle D^m \eta, a_{m-1} D^{m-1} \nu \rangle$ and $\langle a_{m-1} D^{m-1} \eta, D^m \nu \rangle$. We show it for the latter term.

Let $\overline{P} \in S_k^{-1}(\Delta)$ interpolate $a_{m-1} D^m \nu$ at $\{\tau_{ij}\}$, $i = 1, \ldots, N$, $j = 1, \ldots, k$. Then

(5.10) $$\|\overline{P}\|_{L_\infty(I)} \leq O(h^k); \quad \|a_{m-1} D^m \nu - \overline{P}\|_{L_\infty(I)} \leq O(h^k) .$$

Now,

(5.11) $$\langle D^{m-1} \eta, a_{m-1} D^m \nu \rangle = \langle D^{m-1} \eta, \overline{P} \rangle$$

and

(5.12) $$\langle D^{m-1} \eta, \overline{P} \rangle = (D^{m-1} \eta, \overline{P})$$

since on each subinterval I_i, $D^{m-1} \eta \cdot \overline{P} \in \mathcal{P}_{2k}(I_i)$. Hence, we can estimate

(5.13) $$|\langle a_{m-1} D^{m-1} \eta, D^m \nu \rangle| = |(D^{m-1} \eta, \overline{P})| \leq \|D^{m-1} \eta\|_1 \|\overline{P}\|_\infty \leq O(h^{2k+1}) .$$

For the term $\langle D^m \eta, a_{m-1} D^{m-1} \nu \rangle$, the estimate is similar.

We have shown (5.3) and the conclusion (5.8) follows. Q.E.D.

We are now ready for the general results.

Theorem 5.1.

Under our assumptions, we have (1.18) and (1.19):

$$(5.14) \qquad \|D^{\ell} e\|_{L_\infty(I)} \le O(h^{k+\min(n, m-\ell)}), \qquad \ell = 0, \ldots, m$$

and for $0 \le i \le N$

$$(5.15) \qquad |D^{\ell} e(t_i)| \le O(h^{k+n}), \qquad \ell = 0, \ldots, 2(m-1) - q .$$

Proof.

Let $0 \le \ell \le m - 1$. Assume, without loss of generality, that $n \ge m - \ell$. We can write for any $v \in \mathring{S}$

$$(5.16) \qquad D^{\ell} e(t) = (R, G_{\ell}(t)) = [(R, G_{\ell}(t)) - \langle R, G_{\ell}(t)\rangle] +$$

$$+ [\langle R, G_{\ell}(t) + Lv\rangle - (R, G_{\ell}(t) + Lv)] + (R, G_{\ell}(t) + Lv) .$$

The first two terms on the right hand side of (5.16) can be estimated by Theorem 2.2, since we have proven (2.18). We are left with

$$(R, Lv + G_{\ell}(t)) .$$

Now,

$$|(R, Lv + G_{\ell}(t))| \le \|R\|_\infty \|G_{\ell}(t) + Lv\|_1 ,$$

so that we have left to show that

$$(5.17) \qquad \min_{v \in \mathring{S}} \|G_{\ell}(t) + Lv\|_1 \le O(h^{m-\ell}) .$$

Let $t \in I_{i_0}$. We can replace $G_{\ell}(t, \cdot)$ in I_{i_0} by a polynomial piece, say, such that the resulting function ψ will satisfy

$$\|G_{\ell}(t) - \psi\|_{L_\infty(I_{i_0})} \le O(h^{m-\ell-1}) ,$$

so that

$$(5.18)\quad \|G_\ell(t)-\psi\|_{L_1(I)} = \sum_{i=1}^{N} \|G_\ell(t)-\psi\|_{L_1(I_i)} \le h_{i_0} \|G_\ell(t)-\psi\|_{L_\infty(I_{i_0})} \le O(h^{m-\ell})$$

and

$$\|D^{m-\ell}\psi\|_{L_\infty(I_i)} \le \begin{cases} O(1) & i \ne i_0 \\ O(h^{-1}), & i = i_0 . \end{cases}$$

Now, for ψ, clearly there exists a $v \in \mathring{S}$ such that

$$(5.19)\qquad \|Lv-\psi\|_{L_1(I)} \le O(h^{m-\ell}) \|D^{m-\ell}\psi\|_{L_1(I)} \le O(h^{m-\ell}) .$$

Collecting (5.18) and (5.19) we obtain (5.17), hence (5.14), for $0 \le \ell \le m-1$. The result for $\ell = m$ now follows from (5.8).

As in Theorem 3.1, we obtain nothing better for $t = t_i$ and $\ell > 2(m-1)-q$. We now prove the superconvergence result, assuming $n > m-\ell$, with $\ell \le 2(m-1)-q$.

Let now $t = t_i$ for some i. We break the interval $[a,b]$ into two.

In $[a,t_i]$ we have $G_\ell(t_i,\cdot) \in W_\infty^n[a,t_i]$ and in $[t_i,b]$ we have $G_\ell(t_i,\cdot) \in W_\infty^n[t_i,b]$. Also $G_\ell(t_i,\cdot) \in C^{m-\ell-2}[a,b]$ and, since $\ell \le 2(m-1)-q$, we get $m-\ell-2 \ge q-m$, and

$$G_\ell(t_i,\cdot) \in C^{q-m}[a,b] .$$

Let now $\phi(\cdot)$ satisfy (1.2) and

$$(5.20)\qquad L\phi(\cdot) = G_\ell(t_i,\cdot) \text{ on } I .$$

Such ϕ exists by our basic assumption on L and, moreover, we have

$$(5.21)\quad \begin{cases} \phi \in W_{\infty}^{n+m}[a, t_i] \\ \phi \in W_{\infty}^{n+m}[t_i, b] \\ \phi \in C^q[a, b] \,. \end{cases}$$

But then, we can find $v_1 \in \mathring{S}^q_{k+m}(\Delta)$ to "fill the hole" of low continuity of ϕ at t_i, i.e., $\exists v_1 \in \mathring{S}$ such that

$$v_1 + \phi \in W^{n+m}[a, b]$$

(cf. [10]). Now, for $v_1 + \phi$, clearly there exists $\tilde{v} \in \mathring{S}$ such that

$$\| D^j(\tilde{v} - (v_1 + \phi)) \|_{L_\infty(I)} \le O(h^{n+m-j}), \qquad 0 \le j \le m$$

(see e.g. [2]). Denoting $v = v_1 - \tilde{v}$ we obtain

$$\| Lv + G_{\ell}(t_i) \|_{L_\infty(I)} \le O(h^n)$$

and (5.15) follows. Q.E.D.

6. Continuous least squares and nonlinear problems

In this section we extend our results to the integral least squares method and to nonlinear problems; first, the integral (or "continuous") least squares method.

Our effort was concentrated, up to here, on the discrete formulation (1.15), because this is the approximation scheme actually calculated with. Note, though, that by a similar, and shorter, analysis the results are immediately extendible to the continuous least squares method (1.16).

Theorem 6.1.

Let w^* be the solution of (1.16). Under our smoothness and quasiuniformity assumptions, results similar to (1.18), (1.19) are obtained for w^* replacing v^*.

Proof.

We simply go through the much simplified versions of the proofs for the discrete case in Sections 3, 4, 5, and apply them to the continuous case. Q.E.D.

We now consider the extension of our results to the nonlinear ordinary differential equation

$$(6.1) \qquad D^m u(t) = F(t, u(t), Du(t), \dots, D^{m-1}u(t)), \quad a \le t \le b$$

subject to the side conditions (1.2). (If the side conditions were non-homogeneous, a reduction to the homogeneous case is done in a trivial manner). We proceed in the same way as in [5], [17], i.e. apply a

local linearization and iteration process (Newton's method). Since the method of linearization is independent of the numerical method used to approximate the linearized problem, the results of [5], [17] will apply here.

To be able to use any of these numerical methods, the sought solution u of (6.1)-(1.2) needs to be an isolated one. This will happen if the linearized problem at u is uniquely solvable: Consider the curve $\mathbb{C} \subset \mathbb{R}^{m+1}$ defined by

$$\mathbb{C} \equiv \{[t, u(t), \ldots, D^{m-1}u(t)]^T, \ t \in [a,b]\}$$

and suppose that $F \in C^1[\bar{\eta}]$, where $\eta \subset \mathbb{R}^{m+1}$ is some δ-neighborhood of $\mathbb{C}$. Define, for $0 \leq \ell \leq m-1$,

$$(6.2)\quad F_\ell(t;v) = \frac{\partial F}{\partial z_\ell}(t, z_0, \ldots, z_{m-1})\Big|_{z_i = D^i v,\ i=0,\ldots,m-1}, \quad v \in C^{m-1}[a,b]$$

and

$$(6.3)\qquad L = D^m - \sum_{\ell=0}^{m-1} a_\ell(\cdot)D^\ell$$

with

$$(6.4)\qquad a_\ell(t) := F_\ell(t;u), \quad \ell = 0, \ldots, m-1 .$$

Then, if Green's function $G(t,s)$ for L of (6.3) exists, implying unique existence for the linearized problem, this is sufficient to guarantee that $\exists\sigma > 0$ such that u is the unique solution of (6.1)-(1.2) in the sphere $B(D^m u, \sigma)$ [17].

For Newton's method to converge quadratically we have to assume more about F. A sufficient assumption is that $F \in C^2(\overline{\eta})$. The actual process is as follows: a sequence of approximations $\{v_r\}_{r \geq 0}$ is determined with $v_0 \in \mathring{S}$ being a (good) initial guess and $v_{r+1} \in \mathring{S}$ being the discrete least squares solution to

$$(6.5) \qquad \min_{v \in \mathring{S}^q_{k+m}(\Delta)} |L_r v - f_r|_2^2$$

where

$$L_r v(t) = D^m v(t) - \sum_{\ell=0}^{m-1} F_\ell(t; v_r) D^\ell v(t)$$

$$(6.6)$$

$$f_r(t) = F(t, v_r, \ldots, D^{m-1} v_r) - \sum_{\ell=0}^{m-1} F_\ell(t; v_r) D^\ell v_r(t) .$$

For purposes of analysis it is more convenient to consider $y \equiv D^m u$, for which (6.1)-(1.2) are written as

$$(6.7) \qquad y = Ty$$

where

$$(Tf)(t) = F(t, f^{(-m)}(t), \ldots, f^{(-1)}(t))$$

$$(6.8) \qquad f^{(-m+\ell)}(t) = (H_\ell(t), f), \qquad \ell = 0, \ldots, m-1 ,$$

$H_\ell(t, s)$ given by (3.4). The approximation $Y \in S_k^{q-m}(\Delta)$ (to which $\{D^m v_r\}_{r \geq 0}$ will eventually converge) that we wish to find satisfies

$$(6.9) \qquad Y = P_\Delta TY$$

where P_Δ is the discrete L_2-projector; i.e. P_Δ associates with a

function $f \in C[a,b]$ an approximation $P_\Delta f \in S_k^{q-m}(\Delta)$ such that

(6.10) $$|P_\Delta f - f|_2^2 \le |w - f|_2^2, \quad w \in S_k^{q-m}(\Delta).$$

It has been shown in [10], [4] that the L_2-projector is bounded in the uniform norm, provided that the mesh Δ is quasiuniform. The boundedness of the discrete L_2-projector, P_Δ, follows as in Lemma 3.1. Thus we can obtain the following theorem, referring to the analysis in [5] (Thm. 3.1 there; cf. also [21]).

<u>Theorem 6.2.</u>

Let $u \in C^{(m+k+n)}[a,b]$ be an isolated solution for (6.1)-(1.2) and suppose that $F(t, z_0, \ldots, z_{m-1})$ is sufficiently smooth near u. Then, for h small enough, the approximate solution $v^* \in \mathring{S}_{k+m}^q(\Delta)$, solving

(6.11) $$\min_{v \in \mathring{S}_{k+m}^q(\Delta)} |D^m v(\cdot) - F(\cdot, v, \ldots, D^{m-1} v)|_2^2 = |D^m v^*(\cdot) - F(\cdot, v^*, \ldots, D^{m-1} v^*)|_2^2$$

exists and is unique in some neighborhood of u, and Newton's method (6.5)-(6.6), for approximately solving (6.11), converges quadratically in some neighborhood of v^*. We have $D^m v^* = Y$, where Y is defined by (6.9). Moreover, the error estimates (1.18)-(1.19) hold for v^* here.

7. Numerical examples

The collocation method of de Boor - Swartz [5] has become popular enough in recent years, so that we feel it appropriate to use this method as a frame of reference in comparing the discrete least squares schemes.

All the examples here are of second order, $m = 2$, on $I = [0,1]$, with Dirichlet boundary conditions

$$u(0) = u(1) = 0 \tag{7.1}$$

and are smooth enough so that $n = k$.

Unless otherwise specified (as is in example 2), we have approximated using quintics, i.e. $k = 4$, and on a uniform mesh. Thus $h = 1/N$, where N is the number of intervals. We have used a B-spline basis [3] for the approximation space $\mathring{S} = \mathring{S}^q_{k+m}(\Delta)$:

$$\mathring{S} = \text{span}\{\varphi_1, \ldots, \varphi_{\dim}\}\ .$$

We note that all the schemes here evaluate the differential equation at exactly the same points, namely, the Gaussian points $\{\tau_{ij}\}$ of (1.9). The collocation solution [5] is in $C^1[0,1]$; thus it has enough free parameters to satisfy the differential equation at all $\{\tau_{ij}\}$, $i = 1, \ldots, N$, $j = 1, \ldots, k$. The resulting matrix of the linear system in this case has an "almost block diagonal" structure, elaborated in [7], [6]. The method proposed in [7] was used here for the solution of the linear system of algebraic equations resulting from the collocation scheme $C - 1$. When raising the continuity at the knots, by requiring the approximate

solution to be in $C^2[0,1]$, $C^3[0,1]$ or $C^4[0,1]$, we obtain, on trying to collocate on $\{\tau_{ij}\}$, an overdetermined system of linear equations. This system is solved in the weighted least squares sense, as explained in the introduction. We denote the schemes obtained by C-2, C-3 and C-4, respectively. The normal equations were solved using banded Cholesky decomposition. We note that, compared to C-1, the condition number of the matrices in the other schemes has been essentially squared, but in practice we have not sensed any effect of that. This indicates a stable structure obtained when applying B-splines at Gaussian points.

We use a notation similar to [5]: we measure the quantities

$$e^{(j)} = \| D^j(u - v^*) \|_{\ell_\infty(I)}, \qquad j = 0,1,2$$

$$e_\Delta^{(j)} = \max_{t \in \Delta} | D^j(u - v^*)(t) |, \qquad j = 0,1,2$$

and use the shorthand

$$.5 - 4 \quad \text{for} \quad .5 \times 10^{-4} .$$

The global uniform errors $e^{(j)}$ have been measured on an "evaluation grid" consisting of the knots $(\pm 10^{-10})$ and 39 equally spaced points between each two knots. Further, in order to measure the rates of convergence, we have listed in row i, $i > 1$, the numbers

$$\gamma^{(j)} = \log(e_i^{(j)} / e_{i-1}^{(j)}) / \log(N_{i-1}/N_i)$$

and $\gamma_\Delta^{(j)}$, similarly defined, where $e_i^{(j)}$ is the error $e^{(j)}$ and N_i is

the number of (uniform) intervals N in the i-th entry. When listing the results, we have omitted the column $e_{\Delta}^{(2)}$ since it almost always coincided with $e^{(2)}$; i.e., the largest errors in the second derivatives were usually obtained at the knots.

Example 1.

We first compute for the simple example appearing in [17], [5]

$$u'' - 4u = 4\cosh(1) \tag{7.2}$$

whose solution is

$$u(t) = \cosh(2t-1) - \cosh(1)\ . \tag{7.3}$$

Results for the schemes C-1 [5], C-2, C-3 and C-4 are contained in Table 1, where under "N" we have listed the number of uniform subintervals, and under "dim" the dimension of the space $\mathring{S}_{k+m}^{q}(\Delta)$ (which is also the order of the system of equations solved). The actually computed values are

$$\alpha_1, \ldots, \alpha_{\dim}$$

for the representation of the approximate solution

$$v^*(t) = \sum_{i=1}^{\dim} \alpha_i \varphi_i(t)\ . \tag{7.4}$$

The theoretical rates of convergence and superconvergence (1.18)-(1.19) (and also (1.11)-(1.12)) are clearly demonstrated in Table 1. Note also the closeness of the C-2 solution to the collocation solution C-1. They remained close in all the examples we tried.

An operation count for the collocation and discrete least squares methods has been given in [18]. In [6] it has been noted that the collocation method C-1 would always be faster, if the "almost block diagonal" structure of the collocation matrix would be used to advantage (rather than merely treating the matrix as "banded"). Our timing results support the claims in [18], [6]: In all our computations, the time required to set up and solve for a certain number N moderately increased according to the following order from left to right

C-1, C-4, C-3, C-2 .

The more significant measure is, however, the amount of work required to obtain a required accuracy. This measure of computing time would order the schemes (according to all our computations) in the following sequence

C-1, C-2, C-3, C-4

C-1 being the most efficient, and C-4 the least efficient.

The above observations, plus the additional flexibility and better condition number, give collocation superiority over discrete least squares. In fact, the only reason we see for not always preferring collocation is that the storage of the already computed solution gets more expensive as the dimension of the approximation space gets higher. The C-2 solution is quite close to the collocation solution, with less free parameters, and still partly possessing the superconvergence phenomenon. (Note, though, that with collocation we can get superconvergence also for the m-th derivative, using the differential equation and the values of the approximate solution and its first m - 1 derivatives).

Table 1

Errors, convergence rates and computing time for example 1.

	N	dim	$e^{(0)}$	$\gamma^{(0)}$	$e^{(1)}$	$\gamma^{(1)}$	$e^{(2)}$	$\gamma^{(2)}$	$e_{\Delta}^{(0)}$	$\gamma_{\Delta}^{(0)}$	$e_{\Delta}^{(1)}$	$\gamma_{\Delta}^{(1)}$
	2	8	.35-5		.56-4		.29-2		.31-8		.78-7	
	3	12	.33-6	5.8	.81-5	4.8	.61-3	3.8	.11-9	8.3	.33-8	7.8
	4	16	.63-7	5.8	.20-5	4.8	.20-3	3.9	.12-10	7.7	.33-9	7.9
	5	20	.17-7	5.9	.68-6	4.9	.84-4	3.9	.19-11	8.2	.57-10	8.0
C-1	6	24	.58-8	5.9	.28-6	4.9	.41-4	3.9	.46-12	7.8	.13-10	8.0
	7	28	.23-8	5.9	.13-6	4.9	.23-4	3.9	.13-12	8.1	.39-11	8.0
	8	32	.11-8	5.9	.68-7	4.9	.13-4	3.9	.46-13	7.9	.13-11	8.0
	9	36	.53-9	5.9	.38-7	4.9	.84-5	3.9	.18-13	8.1	.52-12	8.0
	10	40	.28-9	5.9	.23-7	4.9	.55-5	3.9	.79-14	7.8	.23-12	8.0
	2	7	.35-5		.56-4		.29-2		.31-8		.78-7	
	3	10	.34-6	5.7	.83-5	4.7	.62-3	3.8	.77-10	9.1	.10-5	-6.4
	4	13	.64-7	5.8	.21-5	4.8	.20-3	3.9	.81-11	7.8	.25-6	5.0
	5	16	.17-7	5.9	.69-6	4.9	.85-4	3.9	.15-11	7.6	.76-7	5.3
C-2	6	19	.60-8	5.9	.28-6	4.9	.42-4	3.9	.39-12	7.5	.28-7	5.5
	7	22	.24-8	5.9	.13-6	4.9	.23-4	3.9	.11-12	7.8	.12-7	5.6
	8	25	.11-8	5.9	.69-7	4.9	.14-4	3.9	.41-13	7.7	.56-8	5.6
	9	28	.54-9	5.9	.39-7	4.9	.85-5	3.9	.18-13	7.1	.29-8	5.7
	10	31	.29-9	5.9	.23-7	4.9	.56-5	3.9	.61-14	10.2	.16-8	5.7
	2	6	.11-4		.16-3		.55-2		.11-4		.58-6	
	3	8	.18-5	4.4	.35-4	3.7	.14-2	3.4	.16-5	4.8	.93-5	-6.8
	4	10	.43-6	5.1	.11-4	4.2	.47-3	3.7	.43-6	4.6	.30-5	4.0
	5	12	.12-6	5.6	.39-5	4.5	.20-3	3.8	.12-6	5.6	.11-5	4.6
C-3	6	14	.45-7	5.6	.16-5	4.7	.99-4	3.9	.45-7	5.6	.45-6	4.8
	7	16	.18-7	5.8	.78-6	4.8	.55-4	3.9	.18-7	5.8	.21-6	4.8
	8	18	.83-8	5.9	.41-6	4.8	.33-4	3.9	.83-8	5.9	.11-6	4.8
	9	20	.42-8	5.9	.23-6	4.8	.21-4	3.9	.42-8	5.9	.64-7	4.8
	10	22	.22-8	5.8	.14-6	4.8	.14-4	3.9	.22-8	5.8	.39-7	4.8
	11	24	.13-8	5.8	.89-7	4.8	.94-5	3.9	.13-8	5.8	.25-7	4.8
	12	26	.78-9	5.8	.58-7	4.8	.67-5	3.9	.78-9	5.8	.16-7	4.8

Table 1 (cont.)

	N	dim	$e^{(0)}$	$\gamma^{(0)}$	$e^{(1)}$	$\gamma^{(1)}$	$e^{(2)}$	$\gamma^{(2)}$	$e_{\Delta}^{(0)}$	$\gamma_{\Delta}^{(0)}$	$e_{\Delta}^{(1)}$	$\gamma_{\Delta}^{(1)}$
	2	5	.11-4		.16-3		.55-2		.11-4		.58-6	
	3	6	.47-5	2.1	.69-4	2.0	.11-2	3.9	.31-5	3.1	.21-4	-8.8
	4	7	.71-6	6.6	.16-4	5.1	.41-3	3.5	.71-6	5.2	.41-5	5.6
	5	8	.26-6	4.6	.60-5	4.4	.16-3	4.2	.17-6	6.5	.20-5	3.2
C-4	6	9	.80-7	6.4	.24-5	5.0	.81-4	3.8	.60-7	5.6	.74-6	5.6
	7	10	.36-7	5.2	.12-5	4.7	.46-4	3.8	.23-7	6.3	.38-6	4.3
	8	11	.16-7	6.1	.62-6	4.8	.27-4	3.9	.11-7	5.5	.20-6	4.8
	9	12	.82-8	5.6	.35-6	4.8	.18-4	3.8	.53-8	6.0	.12-6	4.6
	10	13	.44-8	5.9	.21-6	4.8	.12-4	3.8	.30-8	5.6	.71-7	4.8
	11	14	.25-8	5.8	.13-6	4.8	.81-5	3.8	17-8	5.9	.45-7	4.7
	12	15	.15-8	5.9	.88-7	4.8	.58-5	3.9	.10-8	5.7	.30-7	4.8
	13	16	.95-9	5.8	.60-7	4.8	.43-5	3.9	.64-9	5.8	.20-7	4.8
	14	17	.61-9	5.9	.42-7	4.9	.32-5	3.9	.42-9	5.8	.14-7	4.8
	15	18	.41-9	5.9	.30-7	4.9	.25-5	3.9	.28-9	5.8	.10-7	4.8

Example 2.

Here we consider a less trivial problem of boundary layer, similar to that considered in [17], [20]

$$\varepsilon u'' - (2 - t^2)u = -1 + \frac{1}{2}(2 - t^2)(1 - t) + (1 - t^2)\exp\{-\frac{1 - t}{\sqrt{\varepsilon}}\} \tag{7.5}$$

with a very good approximation to the exact solution, for a small $\varepsilon > 0$, given by

$$u(t) \approx \frac{1}{2 - t^2} - \exp\{-\frac{1 - t}{\sqrt{\varepsilon}}\} + \frac{1}{2}(t - 1) . \tag{7.6}$$

The solution varies very fast between $t = 1 - O(\sqrt{\varepsilon})$ and $t = 1$. We took $\varepsilon = 10^{-8}$. (Note that the "trouble making" term of u' is absent in (7.5). However, it is still of interest to observe the performance of the schemes for this rapidly-varying solution).

We followed [17] by choosing the highly nonuniform partition $(N = 23)$

0, .1, .3, .5, .7, .8, .9, .95, .97, .98, .985, .99, .9925, .995, .996, .997, .9975, .998, .9985, .999, .9995, .9997, .9999, 1.0

and have computed piecewise cubic $(k = 2)$ and quintic $(k = 4)$ approximate solutions by the collocation and discrete least squares schemes on Gaussian points. Results are listed in Table 2. Note that we may already be in the asymptotic range of superconvergence for "quintic C-1" in $e_\Delta^{(0)}$ and $e_\Delta^{(1)}$ and for "quintic C-2" in $e_\Delta^{(0)}$. Note also that the errors listed in Table 2 are absolute, rather than relative, and that we have for $t \approx 1 - \sqrt{\varepsilon}$

$$u = O(1), \; u' = O(1/\sqrt{\varepsilon}), \; u'' = O(1/\varepsilon) .$$

Table 2

Boundary layer problem, example 2 $(N = 23)$ $(\varepsilon = 10^{-8})$

method	dim	$e^{(1)}$	$e^{(2)}$	$e^{(3)}$	$e_{\Delta}^{(0)}$	$e_{\Delta}^{(1)}$
cubic C-1	46	.39-2	.95+2	.67+7	.19-2	.27+2
cubic C-2	24	.69-2	.14+3	.67+7	.31-2	.85+2
quintic C-1	92	.29-4	.12+1	.17+6	.63-5	.62-1
quintic C-2	70	.36-4	.14+1	.13+6	.93-5	.72+0
quintic C-3	48	.15-3	.39+1	.21+6	.11-3	.32+1
quintic C-4	26	.94-3	.18+2	.72+6	.92-3	.93+1

Example 3.

We now turn to a physically interesting nonlinear problem [17], [14], [9]

(7.7) $$u'' + \tau \exp\{u\} = 0, \quad \tau > 0 .$$

The solutions of this problem are given by

(7.8) $$u(t) = -2 \log\{\cosh[(x - \tfrac{1}{2})\tfrac{\theta}{2}]/\cosh[\tfrac{\theta}{4}]\}$$

where

(7.9) $$\theta = \sqrt{2\tau} \cosh(\tfrac{\theta}{4}) .$$

There exists a critical value of τ, τ_c, such that equation (7.9) has two, one or no solutions according to τ being $\tau < \tau_c$, $\tau = \tau_c$ or $\tau > \tau_c$, respectively. The above solutions are all the possible solutions of (7.7)-(7.1) [9]; i.e., beyond τ_c there is no solution to the problem. For the two solutions u_1, u_2 which do exist when $\tau < \tau_c$, we have

(7.10) $$u_1(t) < u_2(t), \quad 0 < t < 1$$

and u_2 is unstable [13]. The numerical solution will converge to u_1.

Clearly, the more we increase τ, $\tau < \tau_c$, the more difficult it becomes to obtain an accurate numerical solution. In order to find the critical value τ_c, let θ_c be the corresponding solution of (7.9). We have

$$\frac{1}{4}\theta_c = \mathrm{ctgh}(\frac{1}{4}\theta_c), \quad (\mathrm{ctgh} = \frac{\cosh}{\sinh})$$

and

$$\tau_c = 8/\sinh^2(\frac{1}{4}\theta_c) .$$

The resulting values are, approximately,

$$\tau_c \approx 3.51383, \quad \theta_c \approx 4.8 .$$

We have computed with two values of τ:

(i) $\tau = 1$ (then for u_1, $\theta \approx 1.5171645990508$)

(ii) $\tau = 3.5$ (then for u_1, $\theta \approx 4.5518536628383$) .

We have applied Newton's method, starting with $v_0 \equiv 0$ and computing v_{r+1} the approximate solution, according to the appropriate scheme C - i, of the linear problem

$$\text{(7.10)} \quad \begin{cases} y'' + (\tau \exp\{v_r\})y = (v_r - 1)\tau \exp\{v_r\} \\ y(0) = y(1) = 0 . \end{cases}$$

The iteration process was stopped when v_{r+1} agreed with v_r to within at least 13 significant decimal digits.

(i) For $\tau = 1$, results are accumulated in Table 3. In all cases it took 3 iterations for the sequence $\{v_r\}$ to converge. The properties of convergence and superconvergence are demonstrated.

Table 2

Boundary layer problem, example 2 $(N = 23)$ $(\varepsilon = 10^{-8})$

method	dim	$e^{(1)}$	$e^{(2)}$	$e^{(3)}$	$e_{\Delta}^{(0)}$	$e_{\Delta}^{(1)}$
cubic C-1	46	.39-2	.95+2	.67+7	.19-2	.27+2
cubic C-2	24	.69-2	.14+3	.67+7	.31-2	.85+2
quintic C-1	92	.29-4	.12+1	.17+6	.63-5	.62-1
quintic C-2	70	.36-4	.14+1	.13+6	.93-5	.72+0
quintic C-3	48	.15-3	.39+1	.21+6	.11-3	.32+1
quintic C-4	26	.94-3	.18+2	.72+6	.92-3	.93+1

Example 3.

We now turn to a physically interesting nonlinear problem [17], [14], [9]

(7.7) $$u'' + \tau \exp\{u\} = 0, \qquad \tau > 0 .$$

The solutions of this problem are given by

(7.8) $$u(t) = -2 \log\{\cosh[(x - \tfrac{1}{2})\tfrac{\theta}{2}]/\cosh[\tfrac{\theta}{4}]\}$$

where

(7.9) $$\theta = \sqrt{2\tau} \cosh(\tfrac{\theta}{4}) .$$

There exists a critical value of τ, τ_c, such that equation (7.9) has two, one or no solutions according to τ being $\tau < \tau_c$, $\tau = \tau_c$ or $\tau > \tau_c$, respectively. The above solutions are all the possible solutions of (7.7)-(7.1) [9]; i.e., beyond τ_c there is no solution to the problem. For the two solutions u_1, u_2 which do exist when $\tau < \tau_c$, we have

(7.10) $$u_1(t) < u_2(t), \qquad 0 < t < 1$$

and u_2 is unstable [13]. The numerical solution will converge to u_1.

(ii) For $\tau = 3.5$, results are listed in Table 4. As expected, the problem is harder here; nevertheless, the schemes still behave well. The number of iterations needed for convergence of the Newton process is considerably higher here, and is listed under the heading "iter".

The above computations were performed in double precision on the Univac 1110 computer at the University of Wisconsin-Madison.

The author wishes to thank Professor C. de Boor for many helpful discussions.

Table 3

Nonlinear example, $\tau = 1$.

	N	dim	$e^{(0)}$	$\gamma^{(0)}$	$e^{(1)}$	$\gamma^{(1)}$	$e^{(2)}$	$\gamma^{(2)}$	$e_\Delta^{(0)}$	$\gamma_\Delta^{(0)}$	$e_\Delta^{(1)}$	$\gamma_\Delta^{(1)}$
	4	16	.43-8		.14-6		.13-4		.65-11		.48-11	
C-1	6	24	.39-9	5.9	.19-7	4.9	.27-5	3.9	.23-12	8.2	.22-12	7.6
	8	32	.71-10	6.0	.45-8	5.0	.87-6	4.0	.16-13	9.4	.48-13	5.3
	4	13	.46-8		.14-6		.14-4		.59-11		.44-7	
C-2	6	19	.40-9	6.0	.19-7	5.0	.28-5	4.0	.22-12	8.1	.43-8	5.7
	8	25	.72-10	6.0	.45-8	5.0	.88-6	4.0	.15-13	9.3	.77-9	5.9
	4	10	.33-7		.83-6		.24-4		.33-7		.29-6	
C-3	6	14	.36-8	5.4	.13-6	4.5	.57-5	3.6	.36-8	5.4	.37-7	5.1
	8	18	.69-9	5.7	.34-7	4.8	.19-5	3.8	.69-9	5.7	.76-8	5.5
	10	22	.19-9	5.8	.12-7	4.9	.80-6	3.9	.19-9	5.8	.22-8	5.7
	4	7	.73-7		.16-5		.41-4		.73-7		.75-6	
	6	9	.47-8	6.8	.17-6	5.6	.68-5	4.4	.47-8	6.8	.71-7	5.8
C-4	8	11	.85-9	6.0	.41-7	5.0	.22-5	4.0	.85-9	6.0	.13-7	5.8
	10	13	.21-9	6.3	.13-7	5.2	.86-6	4.1	.21-9	6.3	.36-8	5.8
	12	15	.70-10	6.0	.51-8	5.0	.42-6	4.0	.70-10	6.0	.12-8	5.8
	14	17	.27-10	6.1	.23-8	5.1	.22-6	4.1	.27-10	6.1	.51-9	5.8

Table 4

Nonlinear example, $\tau = 3.5$.

	N	dim	iter	$e^{(0)}$	$\gamma^{(0)}$	$e^{(1)}$	$\gamma^{(1)}$	$e^{(2)}$	$\gamma^{(2)}$	$e_\Delta^{(0)}$	$\gamma_\Delta^{(0)}$	$e_\Delta^{(1)}$	$\gamma_\Delta^{(1)}$
	4	16	7	.16-5		.53-4		.61-2		.27-6		.56-6	
	6	24	7	.22-6	4.9	.10-4	4.0	.16-2	3.3	.31-8	11.0	.89-8	10.2
C-1	8	32	7	.44-7	5.5	.28-5	4.6	.57-3	3.7	.29-9	8.3	.84-9	8.2
	10	40	7	.12-7	5.7	.98-6	4.7	.24-3	3.8	.50-10	7.9	.14-9	8.2
	12	48	7	.43-8	5.8	.41-6	4.8	.12-3	3.9	.11-10	8.3	.33-10	7.9
	4	13	8	.65-5		.76-4		.71-2		.49-5		.45-4	
	6	19	7	.56-6	6.1	.12-4	4.5	.19-2	3.3	.30-6	6.9	.87-5	4.1
C-2	8	25	7	.68-7	7.3	.32-5	4.7	.62-3	3.8	.18-7	9.8	.16-5	5.8
	10	31	7	.15-7	6.7	.11-5	4.9	.26-3	3.9	.19-8	10.0	.39-6	6.4
	12	37	7	.49-8	6.3	.43-6	4.9	.13-3	4.0	.32-9	10.0	.13-6	6.0
	4	10	8	.72-4		.31-3		.76-2		.58-4		.20-3	
	6	14	8	.74-5	5.6	.52-4	4.4	.24-2	2.8	.50-5	6.1	.19-4	5.8
C-3	8	18	7	.12-5	6.3	.16-4	4.1	.95-3	3.2	.68-6	6.9	.53-5	4.4
	10	22	7	.28-6	6.5	.61-5	4.3	.44-3	3.4	.15-6	6.8	.17-5	5.1
	12	26	7	.85-7	6.6	.27-5	4.5	.23-3	3.6	.44-7	6.7	.73-6	4.6
	4	7	9	.52-3		.18-2		.25-1		.49-3		.17-2	
	6	9	8	.11-3	3.9	.38-3	3.8	.70-2	3.2	.97-4	4.0	.38-3	3.7
	8	11	8	.65-5	9.8	.43-4	7.6	.20-2	4.3	.48-5	10.5	.38-4	8.0
C-4	10	13	7	.87-6	9.0	.12-4	5.6	.76-3	4.3	.57-6	9.6	.74-5	7.3
	12	15	7	.19-6	8.4	.44-5	5.5	.34-3	4.4	.11-6	9.0	.22-5	6.6
	14	17	7	.55-7	7.9	.20-5	5.3	.18-3	4.2	.31-7	8.3	.84-6	6.4

References

1. G. A. Baker, Simplified proofs of error estimates for the least squares method for Dirichlet's problem, Math. Comp. 27 (122) (1973), 229-235.

2. C. de Boor, On uniform approximation by splines, J. Approx. Th. 1 (1968), 219-235.

3. C. de Boor, Package for calculating with B-splines, MRC Technical Summary Report #1333 (1973).

4. C. de Boor, A bound on the L_∞-norm of L_2-approximation by splines in terms of a global mesh ratio, MRC Technical Summary Report #1597 (1975).

5. C. de Boor and B. Swartz, Collocation at Gaussian points, SIAM J. Numer. Anal. 10 (4) (1973), 582-606.

6. C. de Boor and B. Swartz, Comment on "A comparison of global methods for linear two-point boundary value problems", Los Alamos Rep. LA-UR-75-2343 (1975).

7. C. de Boor and R. Weiss, SOLVEBLOK: A package for solving almost block diagonal linear systems with applications to spline approximations and the numerical solution of ordinary differential equations, MRC Technical Summary Report #1625 (1976).

8. J. H. Bramble and A. H. Schatz, Rayleigh-Ritz-Galerkin methods for Dirichlet problem using subspaces without boundary conditions, Comm. Pure Appl. Math. 23 (1970), 653-675.

9. H. T. Davis, Introduction to nonlinear differential and integral equations, New York: Dover (1962).

10. J. Douglas, T. Dupont and L. Wahlbin, Optimal L_∞ error estimates for Galerkin approximations to solutions of two-point boundary value problems, Math. Comp. 29 (130) (1975), 475-483.

11. J. Douglas, T. Dupont and M. F. Wheeler, Some superconvergence results for an H^1-Galerkin procedure for the heat equation, MRC Technical Summary Report #1382 (1973).

12. T. Dupont, A unified theory of superconvergence for Galerkin methods for two-point boundary problems, SIAM J. Numer. Anal. 13 (3) (1976) 362-368.

13. H. Fujita, On the nonlinear equation $\Delta u + e^u = 0$ and $v_t = \Delta v + e^v$, Bull. Amer. Math. Soc. 74 (1969), 132-135.

14. I. M. Gel'fand, Some problems in the theory of quasilinear equations, Uspehi Mat. Nauk 14, 2 (86) (1959), 87-158; English Transl., Amer. Math. Soc. Transl. 29 (1963), 295-381.

15. T. R. Lucas and G. W. Reddien, Some collocation methods for nonlinear boundary value problems, SIAM J. Numer. Anal. 9 (2) (1972), 341-356.

16. F. Natterer, Uniform convergence of Galerkin's method for splines on highly nonuniform meshes, preprint.

17. R. D. Russell and L. F. Shampine, A collocation method for boundary value problems, Numer. Math. 19 (1972), 1-28.

18. R. D. Russell and J. M. Varah, A comparison of global methods for linear two-point boundary value problems, Math. Comp. 29 (132) (1975), 1007-1019.

19. P. Sammon, The discrete least squares method, preprint, Univ. B. C. Vancouver, Canada.

20. J. M. Varah, A comparison of some numerical methods for two-point boundary value problems, <u>Math. Comp.</u> 28 (127) (1974), 743-755.

21. K. A. Wittenbrink, High order projection methods of moment and collocation type for nonlinear boundary value problems, <u>Computing</u> 11 (1973), 255-274.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) MRC-TSR-1654	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) DISCRETE LEAST SQUARES APPROXIMATIONS FOR ORDINARY DIFFERENTIAL EQUATIONS.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Uri Ascher		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709		12. REPORT DATE (11) July 1976
		13. NUMBER OF PAGES 48
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (9) Technical summary rept.,		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report) (12) 52p.

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Discrete least squares
Gaussian points
Collocation

> or =

C to the power q

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The application of the least squares method, using C^q piecewise polynomials of order $k+m$, $k \geq m$, $q \geq m$, for obtaining approximations to an isolated solution of a nonlinear m-th order ordinary differential equation, involves integrals which in practice need to be discretized. Using for this latter purpose the k-point Gaussian quadrature rule in each subinterval, the → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 221 200 LB

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT - Cont'd.

C to the power m-1

cont.

discrete least squares schemes obtained are close to collocation, on the same points, by piecewise polynomials from C^{m-1}.

We prove here that under smoothness assumptions similar to those made by de Boor and Swartz for the collocation procedure, i.e. that the solution be in C^{m+2k}, an optimal global rate of convergence $O(|\Delta|^{k+m})$ is obtained in the uniform norm for the discrete least squares schemes, provided that the partitions Δ are quasiuniform. In addition, a superconvergence rate of $O(|\Delta|^{2k})$ is obtained at the knots for those derivatives ℓ which satisfy $0 \leq \ell \leq 2(m-1) - q$.